DOCUMENT RESUME

ED 095 816 CG 008 966

AUTHOR Cole, Nancy S.
TITLE Unresolved Technical Issues in Fair Interest Measurement.
PUB DATE 74
NOTE 14p.; Paper presented at the Annual Meeting of the American Personnel and Guidance Association (New Orleans, Louisiana, April 1974)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS Discriminatory Attitudes (Social); *Interest Tests; Measurement; Occupational Choice; *Sex Discrimination; Statistical Data; *Test Bias; *Test Validity; *Vocational Interests

ABSTRACT

The problem of sex differences in interest measurement involves many technical issues and procedures. The purpose of this paper is to provide a description of the technical problems involved in construction, scoring, and interpretation of interest measures as related to sex differences and to suggest guidelines within these technical issues which will help eliminate any factors that may influence a person to limit--or cause others to limit--his or her consideration of career solely on the basis of gender. The technical issues revolve around sex differences in item responses, appropriate composition by sex of criterion, reference, or norm groups, what to do about occupations involving primarily one sex, and interpretive and explanatory materials related to the above. Operational guidelines are provided on suitable technical procedures for different types of interest inventory scales as well as helpful interim procedures for use when changes in current procedures require a long period of time for implementation. (Author)

ED 095436

Unresolved Technical Issues in Fair Interest Measurement

Nancy S. Cole
The American College Testing Program

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

This paper is a draft for presentation
at the American Personnel and
Guidance Association Annual Meeting,
New Orleans, April 10, 1974.
The paper should not be cited as a
reference without permission of the author.

CG 008 966

Unresolved Technical Issues in Fair Interest Measurement

Nancy S. Cole
The American College Testing Program

The NIE study of Sex Bias and Sex Fairness in Interest Inventories has been, I believe, a worthwhile effort at defining issues and accomplishing consensus on at least minimal requirements which interest inventories should meet in order to be sex fair. Many issues have been addressed and questions answered through the guidelines resulting from the study. I believe the section of the guidelines on interpretive materials is a very valuable and quite strong section. However, in the technical area the guidelines clearly leave us with some unresolved questions. In this presentation, I will discuss some of the unresolved issues which the guidelines either completely fail to address or address in a neutral form which allows multiple practices. These unresolved areas in the guidelines are a result of incomplete or total lack of information on which to base a reasonable judgment. Consequently I offer these comments not as criticisms, but as suggestions for areas which deserve more careful thought and research in the future so that someday we will be able to resolve the remaining technical issues. In particular, the areas to be discussed are opposite sex occupational scales, item balance within scales, norm groups, and the validity and criterion problem.

Opposite Sex Occupational Scales

Occupational scales of interest inventories have traditionally been constructed upon occupational groups of one sex. For example, if the occupation

of concern is Physician, two scales would be constructed--one on men physicians and one on women physicians. This has been done because there are typically differences in the responses on interest inventory items by men and women physicians and by men and women in general. The side effect of this procedure has been the development of many occupational scales for men for which there are no counterparts for women. Thus one of the earliest goals of those who questioned the fairness of interest inventories was to obtain scores for women on all occupations. The simplest way to get such scores was to score women on the male-construced scales. This has been done on the Kuder for some male-constructed scales for some time and the Strong has recently adopted this procedure. The question remaining is the appropriateness and meaning of scores of members of one sex on scales constructed on the other sex.

The guidelines include elements related to this question but fail to answer the question itself. One guideline states that "scores on all occupations and interest areas covered by the inventory should be given for both males and females." This guideline requires that, at the least, if there is a score for men on some occupation, there must be a corresponding score for that occupation for women. This would be satisfied by a male scale reported to males and a female scale reported to females. Some participants in the NIE study wished the guidelines to require that scores on all scales (male and female) be reported to everyone. However, the feeling of the majority was that the appropriateness of the latter has not been demonstrated. So the possibility of this type of reporting was not excluded but neither was it required.

Because many people desired the stricter requirement that opposite sex scales be reported, let me note the concerns and unanswered questions which have been raised on the issue. The key question is: Do scores on opposite sex scales make any sense? I know of only two types of data which relate to this question. First it has been found that both men and women tend to score higher on opposite sex scales (Campbell, 1974). This suggests a possible lack of meaningfulness of such opposite sex scores. A more important result, reported by Hornaday and Kuder (1961), was that although there are level differences between the sexes on male-constructed scales on the Kuder, the ranking of male-constructed scale scores on women was meaningfully related to a differentiation of occupations. This study showed that, at least for some scales on the Kuder, it is meaningful to report opposite sex scale scores to women as long as those scores are not directly compared with scores on female constructed scales but each set of scales is separately ranked. This result leaves us wondering if a similar ranking procedure within the two sets of scales is possible on the Strong as well. Although results by Campbell (1974) are suggestive that such may not be the case, Campbell's data is on a small and not especially carefully chosen sample of men and women. Thus, we do not yet know the answer to the question of whether reporting scores on opposite sex scales is a meaningful procedure for all inventories or in what particular way such meaningfulness can be accomplished. Another guideline states that "furthermore, reporting of scores for one sex on scales normed or constructed on data from the other sex should be justified by validity data." The effect of both guidelines is to leave it to the inventory

developer to report scores in whatever way the developer can defend by validity data while requiring that both sexes receive scores on the same set of occupations.

Item Balance within Scales

There has been much concern that interest inventory items in some areas are much more closely related to the experience of one sex or another and lead to very different response rates by men and women. An example of this would be a Social item such as "babysitting" or a Mechanical item "fixing a car." Originally the concern was that within the total inventory there seemed to be many more items familiar to men than to women. In recent years a balance of items has been achieved within most widely used interest inventories. However items in some career areas typically remain more familiar to women and in other areas more familiar to men. Some women have argued, "How are we ever going to get any women in mechanical careers, for example, if all the mechanical items in the interest inventory relate to the activities of men and not of women?" They continue, "Women do mechanical things around the house, but inventories never ask about those." The frequent response to such comments is, "We know that the present mechanical items relate to mechanical careers and consequently have validity, but we don't know whether household mechanical items such as 'fixing a sewing machine' would have similar validity." Thus again there is the question of a desired goal but the restriction that we don't yet know if it is technically feasible to meet that goal. Consequently the guideline on this topic is a distinct compromise. The relevant guideline reads "insofar as possible, item pools should tap experiences

and activities which are equally familiar to both males and females. In instances where this is not currently possible, it is essential that, at a minimum, the number of items that are favored by each sex should be balanced. Further it is desirable that the balance of items favored by each sex be achieved within individual scales." The question that remains unresolved is: Will scales constructed with balance within scales have validity? Again we lack the necessary data to answer the question. However, we do know that a substantial portion of the items on present interest inventories are responded to favorably by similar proportions of both sexes. Johansson and Harmon (1972) found that approximately 55-60% of the items common to the male and female forms of the Strong Vocational Interest Blank showed only small sex differences. Thus there is a possible pool of sex balanced items on the Strong, but whether those sex balanced items are the ones which differentiate occupational groups from people in general and therefore could be used in occupational scales is not known. Approximately 40% of the items on the ACT Interest Inventory have response rates for men and women within 10%. However these items are not distributed evenly across the six scales of the inventory. ACT is working with Jack Rayman, a graduate student at the University of Iowa, on a dissertation study to construct sex balanced items and then to examine and attempt to validate them. I believe that at least two other studies of this issue are underway around the country. Hopefully the results will at the least be adequate to answer the question whether or not scales using sex balanced items are technically feasible and further that the studies will demonstrate such feasibility.

Norm Groups

Four concerns involved in a judgment about the best or most appropriate norm group to use on general interest inventory scales are outlined in Table 1. They are the concern with the distribution of scale scores, the form of occupational group profiles, the validity of the scale, list, or map, and the distribution of suggested occupational options. For each concern there are desirable and undesirable outcomes and the unresolved question is whether any particular norming procedure accomplishes all four desirable outcomes.

General scale score distributions. The distribution of general scale scores relates to the previous issue of item balance by scale. Once item responses are combined into a scale score that scale score will reflect the degree of sex balance or sex imbalance in the items. If there is sex balance in the items then the mean score for men and women on the scale will be comparable and the question of norming is not an issue as any norming procedure will provide similar results. If, however, the items are not balanced for the sexes (and consequently the scale scores are sytematically different for the sexes), then the issue of appropriate norm group can be a very important one. If scores are used in this raw scale score form there will be wide sex differences in the distribution of scale scores. Similarly if the scores are used in reference to a combined sex norm group (one composed of equal numbers of men and women), the resulting distribution of scores will be highly imbalanced by sex. However, if the originally imbalanced scores are referenced to the same sex norm group, then the original imbalance will be counterbalanced and

TABLE 1

Concerns in Appropriate Norm Group Selection

#1 Gen'l Scale Score Distributions	#2 Validity of Gen'l Scales, Lists, and Maps	#3 Form of Occupational Profiles	#4 Occupational Option Distributions
A. Balanced by Sex*	A. Valid*	A. Same for Both Sexes*	A. Balanced by Sex*
B. Sex Restrictive	B. Not Valid	B. Different by Sex	B. Sex Restrictive

*In each case, the outcome labeled A is the desirable outcome.

the result will be similar distributions of scale scores by sex. A balanced score distribution by sex has been labeled "non sex restrictive" by Prediger and Hanson (in press) and I believe such a balanced distribution is a desirable goal.

The NIE guidelines require only that the distributions of scores by sex be reported not that such distributions be "non restrictive." In fact, several inventories have general scales which are sex-restrictive as defined. However, since there are multiple concerns in this issue it is possible for a sex-restrictive scale score distribution to lead to more satisfactory outcomes on the other three areas. Thus, we must consider whether all four can be met before censoring an inventory failing to meet one of the concerns.

Form of occupational profiles. Typically, general scale scores are linked to occupations through lists of occupations classified by highest interest scores or through occupational maps. If one is willing to have different occupational classification lists or different occupational maps for men and women, then any norm group can be used. There are, however, two undesirable elements to separate lists or maps by sex. First the appearance of separate treatment may unnecessarily emphasize sex differences in career considerations and may contribute to stereotypes of some occupations being "male" and others "female." Second, if separate maps or lists are necessary, the list or map for each sex will necessarily include only those occupations presently employing members of that sex. Thus, we are back to the old impasse of not having information for women about occupations presently employing only men and vice versa.

From these two problems, it is clear that it is desirable to be able to deal with "occupations" not "male occupation" and "female occupations." To have scale scores in a form which requires only one occupational list or map not only eliminates the possible bad appearance of two lists but has the second important advantage of allowing classifications based on only one sex to serve also as guides for the other sex. If men and women carpenters, plumbers, and engineers have the same or very similar score profiles, it is reasonable to use the score profile for male electricians for both sexes even if presently no female electricians are available.

Here the unresolved issue is: Does the use of a particular norm or reference group contribute to achieving a single occupational map or list for both sexes? Studies of the ACT Interest Inventory (Hanson, in press) showed that through the use of scales normed separately by sex, profiles of educational majors were highly similar. Consequently for this inventory separate sex norms were implemented and a single map of educational majors was used for members of both sexes. However, Holland (Holland et al, 1969) and Campbell (1974) have reported data for occupational groups using raw scores and scores normed on a combined sex group respectively which show some similar and some dissimilar profiles. It is not known whether same sex norming would have achieved uniformly similar profiles in those two cases. Campbell (1974) noted that when different profiles occur to be valid it is essential to use separate occupational lists, and occupations are separately classified in the new manual for the SCII. To use a single occupational classification list (based primarily on men) for both sexes when profiles

differ in many cases as Holland now does in the SDS is not supported by validity data.

Validity of general scale, list, or map. Although some have related the validity of a general scale to the question of norm groups, when the criterion for validity is differentiation of occupational groups it is possible for different norm group procedures, if appropriately applied, to produce equally valid results. Thus, while some type of validity is a necessary condition for an interest inventory, it appears likely to me that valid scores (in the sense of group differentiation) referenced to one norm group could also be valid when using a different norm group. However, when general scale scores are linked to occupations, a new concern for validity is raised. In this use, the validity of an occupational list or map depends upon an accurate reflection of occupational group scores. Thus, as described above if occupational profiles are different by sex, a valid occupational list must reflect those differences. Again, however, the desired goal is a valid single occupational list or map derived from sex balanced general scale scores and, as described next, producing sex balanced occupational options.

Occupational option distribution. Just as there is concern that general scale scores not be restrictive by sex, similarly we wish the occupational options to be non restrictive. That is, the distribution of occupations suggested to people on the basis of their general interest scales should be similar for the two sexes. In some instruments where the occupational options are emphasized this is a more important area for non sex restrictiveness than the general scales. Although we have some data available on the restrictivness

of general scale scores, very little is known about the distribution of occupational options based on the linkage of general scale scores to occupations.

Summary. We have described four desirable characteristics of interest inventory scores which relate to the selection of a norm group. It is an empirical question whether an inventory can meet all four characteristics, and at present we do not know if all four are possible for all inventories. In the future more consideration should be given to the four characteristics in present inventories. At the present time, the study guidelines do not prescribe the use of one type of norm group in preference to others. I believe this is an appropriate stand with present information. More important than the norm group per se is the results of using that norm group. In the future I hope inventories will be able to meet the four desirable characteristics described here. If it is feasible to do so, those characteristics themselves might appropriately be a part of future guidelines.

Validity and the Criterion Problem

The question of validity of an interest inventory and the criteria used to establish that validity pose thorny problems. In one of the background issue papers for the NIE conference, Cole and Hanson (in press) noted that the basic, even if unstated, goal of interest measurement is to suggest careers in which people would be likely to be happy and satisfied. However the typical criterion for interest inventories has been group membership rather than a comprehensive measure of satisfaction, partly because of the availability of group membership information and the difficulty of assessing aspects of job satisfaction. The basic problem with the group membership criterion is that it

reflects the status quo of occupational membership, and it only indirectly addresses the ultimate question of predicting job satisfaction. These limitations could be especially important in the case of women and careers when the status quo results in a distribution of women in a very limited range of careers.

The study guidelines do not address the criterion issue at all. In several places the term "validity" is mentioned but never are any of the many possible types of validity specified. Since group membership has been the primary criterion in the past, one might assume the guidelines require only such criteria in the future. Hopefully, we will begin more careful study of the types of satisfaction people get from careers and the role of inventory interests in relating to some areas of career satisfaction.

References

Campbell, D. P. Manual for the Strong-Campbell Interest Inventory, Stanford, California: Stanford University Press, 1974.

Hornaday, J. A., & Kuder, G. F. A study of male occupational interest scales applied to women. Educational and Psychological Measurement, 1961, 21, 4, 859-864.

Johansson, C. B., & Harmon, L. W. Strong Vocational Interest Blank: One form or two? Journal of Counseling Psychology, 1972, 19, 404-410.

Prediger, D. J. & Hanson, G. R. Definitions of sex-restrictive and sex-biased reporting procedures for interest inventories. Ad hoc paper prepared for the NIE Workshop on Sex Bias and Sex Fairness in Career Interest Inventories, Washington, D. C., March, 1974.

Hanson, G. R. Assessing the Career Interests of College Youth: Summary of Research and Applications. Research Report (in press). Iowa City, Iowa: The American College Testing Program, 1974.

Cole, N. S. & Hanson, G. R. Impact of Interest Inventories on Career Choice. Commissioned paper for the NIE Workshop on Sex Bias and Sex Fairness in Career Interest Inventories, Washington, D. C., March, 1974.